



Configuring, storing, and serving the next-generation global reforecast

Tom Hamill

NOAA Earth System Research Lab

Outline

- Review of DOE grant & constraints
- Principles
- Anticipated users
- Proposed configuration
- Storage benchmarks
- Lists of proposed fields to archive.

What we have been granted by DOE, and what's expected of us

- **14.5 M CPU hours** (CPUh) on “Franklin” supercomputer at Lawrence Berkeley lab, all to be **used before June 2011**.
 - some of those cycles (10%) needed for generating initial conditions, too.
- We are expected to make this data set available to the community by mid-2011 and to generate some experimental forecasts for the renewable energy sector (streamflow, solar potential, wind-energy potential)

Principles

- Reforecasts will be computed with a (smaller-ensemble) version of the GEFS that will be operational in 2011.
- We hope that GEFS will remain in this configuration, or will be changed only slightly, for several years thereafter.
- Once GEFS changes, either EMC or ESRL will continue to run the reforecast version until a next-generation reforecast is in place.

Some anticipated users

- NCEP/CPC, for calibration of probabilistic 6-10 day, week-2 forecasts, development of new longer-lead threats assessments.
- NCEP/HPC and WFOs, for calibration of shorter-range heavy precipitation forecasts.
- OHD, universities, labs, and international organizations for ensemble streamflow calibration and other hydrologic calibration efforts
- Federal, state, local and private water, environmental and energy management agencies; e.g. these data will be used to help operate the NYC water supply system
- NHC, for statistical calibration of track and intensity forecasts of hurricanes, TCgenesis probabilities, etc.
- NCEP/SPC, for longer-lead calibrated severe weather forecast products.
- Universities, labs, ESRL, MDL, EMC, others for development of advanced statistical post-processing techniques, high-impact events.
- ESRL, NCEP/EMC, and universities, for data assimilation research into correcting bias in short-term forecasts.
- ESRL, as advertised in proposal, for R&D of longer-lead forecasts of solar-energy potential, wind-power potential.
- ESRL and EMC, for determining optimal reforecast configuration into the future.
- Many unforeseen users, if we make it easy for them to get the data.

The agreed-upon reforecast configuration is not likely to be optimal for any one of these users, but we hope it will be satisfactory for most to make substantial progress.

A proposed configuration

(based on 16 July 2010 telecon discussion)

- Assume T254L42 GFS model used for 0-8 days. From days 7.5-16, T190L42. Estimated 92.19 CPUh for 16-day forecast on DOE Franklin. Say 30 years, 365 days/year, 10 members, 2x daily \sim 20.2M CPU h , which is greater than 14.5 M total allocation.
- Also, will need perhaps 10% of allocation for generating initial conditions, testing.
- Proposed configuration based on discussion: At 00Z, full 10-member forecast every day for 30 years out to 16 days. At 12Z, a T254L42 ensemble to day 8, 5 members, reforecast every other day. Total = $10.1 + 1.75\text{M CPUh} = 11.85\text{ M CPUh}$. Doable.
- Extra CPU time may be devoted to running some 45-day reforecasts at T126 for the sake of comparison against CFSR reforecasts, with coupled ocean model (it'd be nice to have a few spare cycles to do this).

Benchmarking the storage of full model

- Assume we want a permanent archive of the full model fields. Assume this data infrequently accessed, though if storage solution permits rapid access, great.
- The full output of a T254L42 sigma (upper-air) file is ~64 MB, and the “sflx” (surface) file is ~40 MB. “Restart” file at end of T254L28 run is ~168MB. Full output of a T190L42 sigma file is ~35 MB, and sflx file is 23 MB. “Restart” file at end of T190L28 run is ~93MB.
- Storage for a single 0- to 8-day T254L42 forecast:
 - Output saved 8x daily for 0-3 days, 4x daily for 3-16 days. Restart file at end of day 7.5 for the beginning of the T190L42 forecasts.
 - $[3 \text{ days} * 8x \text{ daily} + 1 + 5 \text{ days} * 4x \text{ daily}] * [64 + 40 \text{ MB}] + 168 \text{ MB restart} = 4.9 \text{ GB storage}$
- Storage for a single 7.5-16 day T190L42 forecast:
 - Output saved 4x daily for 7.5 to 16 days, plus restart file at end of day 16.
 - $[8.5 \text{ days} * 4x \text{ daily} + 1] * [35 + 23 \text{ MB}] + 93 \text{ MB} = 2.2 \text{ GB}$
- 00Z run to 16 days: $10 \text{ mbrs} * 365 \text{ days} * (4.9 + 2.2 \text{ GB}) * 30 \text{ years} \sim 778 \text{ TB}$
- 12Z run to 8 days: $5 \text{ mbrs} * (365/2 \text{ days}) * 4.9 \text{ GB} * 30 \text{ years} \sim 134 \text{ TB}$
- **Total $\sim 912 \text{ TB}$ for permanent archive of full fields**

Benchmarking, part 2:

user-selected fields on fast access.

- Assume users will want some subset of fields on fast-access. Fields are ~ those stored in TIGGE and NAEFS data sets, plus a few extra. See following slides.
- For selected fields and T254 on 768*384 grid, ~13 MB per time in Grib-2.
- For selected fields and T190 on 576*288 grid, ~7.3 MB per time in Grib-2.
- Storage of one run at 00Z cycle: For 0-3 days, 8x daily, 3-8 days, 4x daily = 45 output times. For days 7.5-16 = 8.5 days * 4x daily + 1 = 35 output times: So $45 * 13 \text{ MB} + 35 * 7.3 = 841 \text{ MB}$.
- Storage of one run to 8 days at 12Z cycle: $45 \text{ output times} * 13 \text{ MB} = 585 \text{ MB}$
- Storage of proposed 30-year reforecast for 00Z configuration = $841 \text{ MB} * 30 \text{ years} * 365 \text{ days} * 10 \text{ members} = 92 \text{ TB}$.
- Storage of proposed 30-year reforecast for 12Z configuration = $585 \text{ MB} * 30 \text{ years} * (365/2 \text{ days}) * 5 \text{ members} = 17 \text{ TB}$.
- **92 + 17 = 109 TB for full TIGGE data set. Also, will need 2x for backup, so 218 TB.**

To make this useful to you, we need ...

- Fast-access storage for $\sim 10^9$ TB data, with redundancy
- Mass storage for ~ 912 TB data, ideally with redundancy.
- Servers, cooling for servers.
- Software to make it easy for you to fetch the data you need.
- Without an “enterprise” solution, costs will scale approximately linearly, so storage costs for 5 members will be roughly half that for 10 members.
- Lesson: **we need to really have users for the data that we store.** We have only \$175K of committed funds now available for storage costs.
 - If you are urging we store even more, can you help pay for the extra archival costs?

Proposed fields for “fast” archive

- Mean and every member
- “Fast” archive will be on disk, readily accessible (as opposed to full model output that may be on some slower archival system)
- Mandatory level data: Geopotential height, temperature, u, v, specific humidity at 1000, 925, 850, 700, 500, 300, 250, 200, 100, 50, and 10 hPa.
- PV ($\text{K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$) on $\theta = 320\text{K}$ surface.
- Wind components, potential temperature on 2 PVU surface.
- Fixed fields saved once:
 - field capacity
 - wilting point
 - land-sea mask
 - terrain height

blue = new fields based on recent discussions

Proposed single-level fields for “fast” archive

- Surface pressure (Pa)
- Sea-level pressure (Pa)
- Surface (2-m) temperature (K)
- Skin temperature (K)
- Maximum temperature since last storage time (K)
- Minimum temperature since last storage time (K)
- Soil temperature (0-10 cm; K)
- Volumetric soil moisture content (proportion, 0-10 cm) – kg/m³ in TIGGE archive
- Total accumulated precipitation since beginning of integration (kg/m²)
- Precipitable water (kg/m², vapor only, no condensate)
- Specific humidity at 2-m AGL (kg/kg; instantaneous) – surface dewpoint in TIGGE archive
- Water equivalent of accumulated snow depth (kg/m²) – from beginning of integration?
- CAPE (J/kg)
- CIN (J/kg)
- Total cloud cover (%)
- 10-m u- and v-wind component (m/s)
- 80-m u- and v-wind component (m/s)
- Sunshine duration (min)
- Snow depth water equivalent (kg/m²)
- Runoff
- Solid precipitation
- Liquid precipitation
- Vertical velocity (850 hPa)
- Geopotential height of surface
- Wind power (=windspeed³ at 80 m*density)
- **Saturation vapor pressure deficit**

Proposed fields for “fast” archive

- Fluxes (W/m^2 ; average since last archive time)
 - sensible heat net flux at surface
 - latent heat net flux at surface
 - downward long-wave radiation flux at surface
 - upward long-wave radiation flux at surface
 - upward short-wave radiation at surface
 - downward short-wave radiation flux at surface
 - upward long-wave radiation at nominal top
 - ground heat flux.